Journal of the Chinese
Statistical Association
Vol. 47, (2009) 104–112

COMPATIBLE NON-NORMALIZED CONDITIONAL DENSITIES

Kun-Lin Kuo

Institute of Statistical Science, Academia Sinica

ABSTRACT

Checking compatibility of the specified conditional distributions is an important problem in statistics, especially in Bayesian computations. However, conditional density may be given without exact normalizing constant multiplier since this multiplier is hard to identify. In this article, we provide necessary and sufficient conditions for compatibility of non-normalized conditional densities. In addition, if they are compatible, we also discuss the uniqueness of the associated joint density which generates them.

Key words and phrases: Compatibility, non-normalized conditionals, uniqueness.
JEL classification: C46.

1. Introduction

The members of a collection of conditional densities are called compatible if there exists a joint density which generates them. For example, suppose that we are given two conditional densities, say $\pi_{X|Y}(x|y)$ and $\pi_{Y|X}(y|x)$, satisfy $\pi_{X|Y}(x|y) > 0$ for all $x \in S_X^y$ and for each $y \in S_Y$, $\pi_{Y|X}(y|x) > 0$ for all $y \in S_Y^x$ and for each $x \in S_X$, and

$$\int_{S_X^y} \pi_{X|Y}(x|y)\, dx = 1, \ \int_{S_Y^x} \pi_{Y|X}(y|x)\, dy = 1.$$

Here, $\bigcup_{y \in S_Y} S_X^y = \bigcup_{x \in S_X} S_Y^x$ and we denote it by S_{XY}. We say that $\pi_{X|Y}$ and $\pi_{Y|X}$

are compatible if there exists a joint density $\pi_{XY}(x,y)$ with support S_{XY} such that

$$\pi_{X|Y}(x|y) = \frac{\pi_{XY}(x,y)}{\int_{S_X^y} \pi_{XY}(t,y)\,dt} \text{ for all } x \in S_X^y \text{ and for each } y \in S_Y,$$

and

$$\pi_{Y|X}(y|x) = \frac{\pi_{XY}(x,y)}{\int_{S_Y^x} \pi_{XY}(x,t)\,dt} \text{ for all } y \in S_Y^x \text{ and for each } x \in S_X.$$

It is well-known that two marginal distributions cannot uniquely determine the joint distribution. However, in specifying multivariate models it is sometimes easier to visualize conditional distributions rather than marginal or joint distributions. For example, a physician would have no difficulty in citing the risk of stoke conditioned on combinations of risk factors such as smoking, obesity, diabetics etc (Ip and Wang, 2009). Applications may be founded in the area of model building in classical statistical settings and in the elicitation and construction of multi-parameter prior distributions in Bayesian scenarios. In addition, spatial model (Besag, 1974), construction of multivariate distributions (Sarabia and Gómez-Déniz, 2008), multiple imputation (Van Buuren et al., 2006; Van Buuren, 2007), and Gibbs sampling with improper posteriors (Hobert and Casella, 1998), are another important areas involving conditionally specified distributions. Unfortunately, there is no guarantee that all conditional models are compatible. Each conditional distribution only contains partial information about the target (joint) distribution. Combining these conditional distributions, one may derive possible conflicting information about the target distribution. Therefore, the compatibility issue is a basic framework for topics related to conditionally specified distributions.

Three common issues of compatibility are (1) whether the given conditional densities are compatible; (2) if they are, whether or not the associated joint density is unique; and (3) how to determine all possible joint densities efficiently. These issues have been of fundamental importance in Bayesian computations and in distribution theory, and have been studied extensively in the literature. Some earlier results can be seen in Besag (1974), Arnold and Press (1989), Gupta and Varga (1990), Casella and George (1992), Gelman and Speed (1993, 1999), Ng (1997), Hobert and Casella (1998), Arnold et al. (1999, 2001, 2002, 2004), Pérez-Villalta (2000), Kopciuszewski

(2004), and Slavkovic and Sullivant (2006). Recently, Wang and Ip (2008) and Ip and Wang (2009) investigated the log-linear interactions to study conditional modelling on a product measurable space (Tian and Tan, 2003). Song et al. (2009) and Tian et al. (2009) gave different methods, respectively, for checking the existence and uniqueness of the joint distribution which is connected with the given finite discrete conditional distributions. Most of related works on conditional models in the literature have focused on conditional density which is given with exact normalizing constant multiplier. However, in practice, the constant multiplier is typically unavailable in closed form and we always take the conditional density to be proportional to a nonnegative function. In this article, we consider non-normalized conditional models and provide methods, respectively, for checking compatibility and uniqueness.

The article is structured as follows. In Section 2, we discuss necessary and sufficient conditions for existence and uniqueness of the two-dimensional joint distribution. Examples are given for illustrating the theory. In Section 3, we provide higher-dimensional results which are analogous to those presented in Section 2. Conclusions are given in Section 4.

2. Two-dimensional cases

Suppose that we are given the following conditional model:

$$\begin{cases} \pi_{X|Y}(x|y) \propto g_1(x,y), \ x \in S_X^y, \ y \in S_Y, \\ \pi_{Y|X}(y|x) \propto g_2(x,y), \ y \in S_Y^x, \ x \in S_X, \end{cases}$$

where $\int_{S_X^y} g_1(x,y)\,dx \neq 0$ for all $y \in S_Y$ and $\int_{S_Y^x} g_2(x,y)\,dy \neq 0$ for all $x \in S_X$. We say g_1 and g_2 are compatible if there exist a joint density $\pi_{XY}(x,y)$, and two functions $c_1(y)$ and $c_2(x)$ such that

$$c_1(y)g_1(x,y) = \frac{\pi_{XY}(x,y)}{\int_{S_X^y} \pi_{XY}(x,y)\,dx} \text{ and } c_2(x)g_2(x,y) = \frac{\pi_{XY}(x,y)}{\int_{S_Y^x} \pi_{XY}(x,y)\,dy}.$$

In the following, we provide a method for checking compatibility of g_1 and g_2.

Theorem 1. *The specified g_1 and g_2 are compatible if and only if there exist two* *functions* $u(x)$ *and* $v(y)$ *such that*

$$\frac{g_1(x,y)}{g_2(x,y)} = \frac{u(x)}{v(y)}, \text{ for all } (x,y) \in S_{XY},$$

where $\iint_{S_{XY}} v(y)g_1(x,y)\,dx\,dy < \infty$.

Proof. Suppose that g_1 and g_2 are compatible and π_{XY} is an associated joint density. Then there exist $c_1(y)$ and $c_2(x)$ such that $c_1(y)g_1(x,y) = \pi_{XY}(x,y)/\pi_Y(y)$ and $c_2(x)g_2(x,y) = \pi_{XY}(x,y)/\pi_X(x)$ where π_X and π_Y are marginal densities of π_{XY}. Then the necessary conditions of the theorem hold.

Conversely, define a joint density as

$$h_{XY}(x,y) = \frac{v(y)g_1(x,y)}{\iint_{S_{XY}} v(y)g_1(x,y)\,dx\,dy}, \ (x,y) \in S_{XY}.$$

Then the corresponding conditional densities are

$$h_{X|Y}(x|y) = \frac{v(y)g_1(x,y)}{\int_{S_X^y} v(y)g_1(x,y)\,dx} \propto g_1(x,y),$$

and

$$h_{Y|X}(y|x) = \frac{v(y)g_1(x,y)}{\int_{S_Y^x} v(y)g_1(x,y)\,dy} = \frac{u(x)g_2(x,y)}{\int_{S_Y^x} u(x)g_2(x,y)\,dy} \propto g_2(x,y).$$

Therefore, g_1 and g_2 are compatible. □

Under the assumptions of Theorem 1 and suppose that g_1 and g_2 are compatible, then an associated joint density can be formulated by

$$v(y)g_1(x,y)\left(\iint_{S_{XY}} v(y)g_1(x,y)\,dx\,dy\right)^{-1}.$$

On the other hand, when g_1 and g_2 are conditional densities, then Theorem 1 is consistent with that given by Arnold and Press (1989, Theorem 4.1).

Example 1. (An incompatible case) Suppose that

$$\pi_{X|Y}(x|y) \propto e^{-xy}, \ x > 0, \text{ and } \pi_{Y|X}(y|x) \propto e^{-xy}, \ y > 0.$$

Here $g_1(x,y) = g_2(x,y) = e^{-xy}$ and we have $g_1(x,y)/g_2(x,y) = 1$. Let $u(x) = c$ and $v(y) = c$ where $c \neq 0$. However, $v(y)g_1(x,y)$ is not integrable over $(0,\infty)\times(0,\infty)$. So, there is no proper joint density with this conditional model.

The next example is given by Kopciuszewski (2004) and we resolved it by using our method.

Example 2. (A compatible case) Suppose that $\pi_{X|Y}(x|y)$ is proportional to the exponential density on the interval $[y-1, y]$, and $\pi_{Y|X}$ is proportional to the normal density, with mean $-x+1/2$ and variance 1, on the interval $[x, x+1]$. That is,

$$\pi_{X|Y}(x|y) \propto g_1(x,y) = e^{-xy}, \ \pi_{Y|X}(y|x) \propto g_2(x,y) = \exp\left(\frac{-(y+x-1/2)^2}{2}\right)$$

and $S_{XY} = \{(x,y) \mid 0 \leq y-x \leq 1, x \geq 0\}$. Notice that

$$\frac{g_1(x,y)}{g_2(x,y)} = \exp\left(\frac{x^2+y^2-x-y+1/4}{2}\right).$$

Set $v(y) = \exp(-y^2/2 + y/2)$. Since $v(y)g_1(x,y)$ is integrable over S_{XY}, we conclude that the conditional model is compatible.

Once we have determined that the specified conditional model is compatible we need to address the issue of whether there is a unique associated joint density. Let V be the collection of all $v(y)$'s in the ratio $g_1(x,y)/g_2(x,y) = u(x)/v(y)$ derived in Theorem 1. We shall, however, consider $v_1(y)$ and $v_2(y)$ to be equivalent if $v_1(y) = cv_2(y)$ almost everywhere for a certain $c \neq 0$. Hence, for all $v_1(y), v_2(y) \in V$, we have $v_1(y) \neq cv_2(y)$ almost everywhere for all $c \neq 0$. Let F consist of all possible joint densities which generate g_1 and g_2. In the following theorem, we show that V and F have the same cardinal number.

Theorem 2. *Suppose that g_1 and g_2 are compatible. Define a mapping $H : V \to F$ by*

$$H(v(y)) = v(y)g_1(x,y)\left(\iint_{S_{XY}} v(y)g_1(x,y)\,dx\,dy\right)^{-1}.$$

Then H is bijective.

Proof. Notice that H is well-defined. Suppose that $\pi_{XY} \in F$, then there exist functions $c_1(y)$ and $c_2(x)$ such that $c_1(y)g_1(x,y) = \pi_{XY}(x,y)/\pi_Y(y)$ and $c_2(x)g_2(x,y) = \pi_{XY}(x,y)/\pi_X(x)$. Obviously, $v(y) \equiv c_1(y)\pi_Y(y) \in V$ and then H is surjective.

Assume that $H(v_1(y)) = H(v_2(y))$. This implies that $v_1(y) = cv_2(y)$ almost everywhere. Therefore, v_1 and v_2 are equivalent and H is then injective. ∎

From Theorem 2, we see that the associated joint density is not unique if there exist non-equivalent v_1 and v_2 such that $g_1(x,y)/g_2(x,y) = u_1(x)/v_1(y) = u_2(x)/v_2(y)$. In addition, all possible joint densities can be constructed by knowing all $v(y)$'s.

Example 3. (A non-unique case) Define the sets $A_1 = \{(x,y) \mid -1 < x, y < 0\}$ and $A_2 = \{(x,y) \mid 0 < x, y < 1\}$ and set $g_1(x,y) = g_2(x,y) = 1$ for $(x,y) \in A = A_1 \cup A_2$.

Clearly, g_1 and g_2 are compatible. Let

$$v_\lambda(y) = \begin{cases} \lambda_1, & y \in (-1,0), \\ \lambda_2, & y \in (0,1), \end{cases}$$

where $\lambda = (\lambda_1, \lambda_2)$ is any positive vector. Therefore, all possible joint densities can be presented by

$$\begin{aligned} & v_\lambda(y) g_1(x,y) \left(\iint_A v_\lambda(y) g_1(x,y)\, dx\, dy \right)^{-1} \\ &= \frac{\lambda_1}{\lambda_1+\lambda_2} I((x,y) \in A_1) + \frac{\lambda_2}{\lambda_1+\lambda_2} I((x,y) \in A_2). \end{aligned}$$

3. Extension to higher dimensions

Results analogous to those presented in the previous section can be given in higher-dimensional settings. Suppose that X_i, $1 \le i \le I$, are random variables with support S_i, respectively. Let $N = \{1, \dots, I\}$, $\alpha \ne N$ be a nonempty subset of N, and $\overline{\alpha} = N - \alpha$. We write the joint density by $\pi_N(\boldsymbol{x}_N)$ on S_N, the α-marginal density by $\pi_\alpha(\boldsymbol{x}_\alpha)$ on S_α, and the conditional density by letting $\pi_{\overline{\alpha}|\alpha}(\boldsymbol{x}_{\overline{\alpha}}|\boldsymbol{x}_\alpha) = \pi_N(\boldsymbol{x}_N)/\pi_\alpha(\boldsymbol{x}_\alpha)$ on $S_\alpha^{\boldsymbol{x}_\alpha}$. Now, we consider the following conditional model:

$$\pi_{\overline{\alpha}_j|\alpha_j}(\boldsymbol{x}_{\overline{\alpha}_j}|\boldsymbol{x}_{\alpha_j}) \propto g_j(\boldsymbol{x}_N), \; 1 \le j \le m.$$

Theorem 3. *The g_j, $1 \le j \le m$, are compatible if and only if there exist $v(\boldsymbol{x}_{\alpha_1})$ and $u_j(\boldsymbol{x}_{\alpha_j})$ for $2 \le j \le m$ such that*

$$\frac{g_1(\boldsymbol{x}_N)}{g_j(\boldsymbol{x}_N)} = \frac{u_j(\boldsymbol{x}_{\alpha_j})}{v(\boldsymbol{x}_{\alpha_1})}, \; \textit{for all } \boldsymbol{x}_N \in S_N,$$

where $\int_{S_N} v(\boldsymbol{x}_{\alpha_1}) g_1(\boldsymbol{x}_N)\, d\boldsymbol{x}_N < \infty$. *If* g_j*'s are compatible, then the associated joint density is not unique if and only if there exist two non-equivalent* $v^{(1)}(\boldsymbol{x}_{\alpha_1})$ *and* $v^{(2)}(\boldsymbol{x}_{\alpha_1})$ *such that*

$$\frac{g_1(\boldsymbol{x}_N)}{g_j(\boldsymbol{x}_N)} = \frac{u_j^{(1)}(\boldsymbol{x}_{\alpha_j})}{v^{(1)}(\boldsymbol{x}_{\alpha_1})} = \frac{u_j^{(2)}(\boldsymbol{x}_{\alpha_j})}{v^{(2)}(\boldsymbol{x}_{\alpha_1})} \text{ for some } 2 \le j \le m.$$

4. Conclusions

The compatibility and uniqueness problems affect the using of the Gibbs sampler and some Markov Chain Monte Carlo methods. In this article, we provide methods for addressing these problems and these methods are very easy to practice. Gelman and Raghunathan (2001) stated that "the study of conditional distributions is an area where theory has not caught up with practice." Our theoretical work here might contribute to some of the compatibility issues.

References

Arnold, B. C., Castillo, E., and Sarabia, J. M. (1999). *Conditional Specification of Statistical Models*, Springer, New York.

Arnold, B. C., Castillo, E., and Sarabia, J. M. (2001). Conditionally specified distributions: An introduction (with discussions). *Statistical Science*, **16**, 249–274.

Arnold, B. C., Castillo, E., and Sarabia, J. M. (2002). Exact and near compatibility of discrete conditional distributions. *Computational Statistics and Data Analysis*, **40**, 231–252.

Arnold, B. C., Castillo, E., and Sarabia, J. M. (2004). Compatibility of partial or complete conditional probability specifications. *Journal of Statistical Planning and Inference*, **123**, 133–159.

Arnold, B. C. and Press, S. J. (1989). Compatible conditional distributions. *Journal of the American Statistical Association*, **84**, 152–156.

Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems (with discussions). *Journal of the Royal Statistical Society. Series B*, **36**, 192–236.

Casella, G. and George, E. I. (1992). Explaining the Gibbs sampler. *The American Statistician*, **46**, 167–174.

Gelman, A. and Raghunathan, T. E. (2001). Comment of Conditionally specified distributions: An introduction. *Statistical Science*, **16**, 268–269.

Gelman, A. and Speed, T. P. (1993). Characterizing a joint probability distribution by conditionals. *Journal of the Royal Statistical Society. Series B*, **55**, 185–188.

Gelman, A. and Speed, T. P. (1999). Corrigendum: Characterizing a joint probability distribution by conditionals. *Journal of the Royal Statistical Society. Series B*, **61**, 483.

Gupta, A. K. and Varga, T. (1990). Characterization of joint density by conditional densities. *Communications in Statistics — Theory and Methods*, **19**, 4643–4652.

Hobert, J. P. and Casella, G. (1998). Functional compatibility, Markov chains, and Gibbs sampling with improper posteriors. *Journal of Computational and Graphical Statistics*, **7**, 42–60.

Ip, E. H. and Wang, Y. J. (2009). Canonical representation of conditionally specified multivariate discrete distributions. *Journal of Multivariate Analysis*, **100**, 1282–1290.

Kopciuszewski, P. (2004). An extension of the factorization theorem to the non-positive case. *Journal of Multivariate Analysis*, **88**, 118–130.

Ng, K. W. (1997). Inversion of Bayes Formula: Explict Formulae for Unconditional pdf, in: N. L. Johnson and N. Balakrishnan, eds. *Advance in the Theory and Practice in Statistics*, Wiley, New York, 571–584.

Pérez-Villalta, R. (2000). Variables finitas condicionalmente especificadas. *Qüestiió*, **24**, 425–448.

Sarabia, J. M. and Gómez-Déniz, E. (2008). Construction of multivariate distributions: a review of some recent results. *SORT*, **32**, 3–36.

Slavkovic, A. B. and Sullivant, S. (2006). The space of compatible full conditionals is a unimodular toric variety. *Journal of Symbolic Computation*, **41**, 196–209.

Song, C.-C., Li, L.-A., Chen, C.-H., Jiang, T. J., and Kuo, K.-L. (2009). Compatibility of finite discrete conditional distributions. To appear in *Statistica Sinica*.

Tian, G.-L. and Tan, M. (2003). Exact statistical solutions using the inverse Bayes formulae. *Statistics and Probability Letters*, **62**, 305–315.

Tian, G.-L., Tan, M., Ng, K. W., and Tang, M.-L. (2009). A unified method for checking compatibility and uniqueness for finite discrete conditional distributions. *Communications in Statistics — Theory and Methods*, **38**, 115–129.

Van Buuren, S., Brand, J. P. L., Groothuis-Oudshoorn, C. G. M., and Rubin, D. B. (2006). Fully conditional specification in multivariate imputation. *Journal of Statistical Computation and Simulation*, **76**, 1049–1064.

Van Buuren, S. (2007). Multiple imputation of discrete and continuous data by fully conditional specification. *Statistical Methods in Medical Research*, **16**, 219–242.

Wang, Y. J., and Ip, E. H. (2008). Conditionally specified continuous distributions. *Biometrika*, **95**, 735–746.

[Received March 2009; accepted July 2009.]